

Riemannian Newton's method for joint diagonalization on the Stiefel manifold with application to ICA

Hiroiyuki Sato*

Department of Applied Mathematics and Physics
Kyoto University, Kyoto 606-8501, Japan

March 31, 2014

Abstract

Joint approximate diagonalization of non-commuting symmetric matrices is an important process in independent component analysis. It is known that this problem can be formulated as an optimization problem on the Stiefel manifold. Riemannian optimization techniques can be used to solve this optimization problem. Among the available techniques, we have studied Riemannian Newton's method for the joint diagonalization problem. In particular, we have shown that the resultant Newton's equation can be effectively solved using the Kronecker product and vec operator. We have performed numerical experiments to show that the proposed method improves the accuracy of an approximate solution of the problem, and have applied the method to independent component analysis.

Keywords: Joint diagonalization; Riemannian optimization; Newton's method; Stiefel manifold; Independent component analysis

1 Introduction

The joint diagonalization (JD) problem for N real $n \times n$ symmetric matrices A_1, A_2, \dots, A_N is often considered on the orthogonal group $O(n)$. The problem is to find an $n \times n$ orthogonal matrix X that minimizes the sum of the squared off-diagonal elements, or equivalently, maximizes the sum of the squared diagonal elements of $X^T A_l X$, $l = 1, \dots, N$ [18]. For more information regarding finding non-orthogonal matrices, see [20]. A solution to the JD problem is valuable for independent component analysis (ICA) and the blind source separation problem [2, 5, 6, 8, 17].

Several approaches have been proposed in the context of Jacobi methods [3, 5, 6] and Riemannian optimization [17, 19]. In [17], the JD problem was considered on the Stiefel manifold $\text{St}(p, n) := \{Y \in \mathbb{R}^{n \times p} \mid Y^T Y = I_p\}$ with $p \leq n$. That is, the required matrix was a rectangular orthonormal matrix. The orthogonal group $O(n)$ is a special case of the Stiefel manifold because $O(n) = \text{St}(n, n)$.

Riemannian optimization refers to optimization on Riemannian manifolds. Unconstrained optimization methods in Euclidean space such as steepest descent, conjugate gradient, and

*hsato@amp.i.kyoto-u.ac.jp

Newton's methods have been generalized to those on a Riemannian manifold [1, 9, 12, 16]. In [17], the Riemannian trust-region method was applied to the JD problem on the Stiefel manifold $\text{St}(p, n)$. With Y varying on $\text{St}(p, n)$ with $p < n$, minimizing the sum of the squared off-diagonal elements of $Y^T A_l Y$, $l = 1, \dots, N$ is no longer equivalent to maximizing the sum of the squared diagonal elements. According to [17], the JD problem on the Stiefel manifold maximizes the sum of the squared diagonal elements of $Y^T A_l Y$, $l = 1, \dots, N$ with $Y \in \text{St}(p, n)$.

This article deals with Newton's method for the JD problem on the Stiefel manifold. The Hessian of the objective function is fundamental for deriving Newton's equation, the key equation in Newton's method. We have intensively examined the Hessian to efficiently solve Newton's equation. In particular, we have used the Kronecker product, and vec and vech operators to transform Newton's equation into the form $Ax = b$.

This paper is organized as follows. We introduce the JD problem on the Stiefel manifold in Section 2, following on from [17]. We also present a brief review of computing the gradient and the Hessian of the objective function. In Section 3, we consider Newton's equation, which is directly obtained by substituting the gradient and Hessian formulas in Section 2 into $\text{Hess } f(Y)[\xi] = -\text{grad } f(Y)$. To derive the representation matrix formula of the Hessian of the objective function, we use the Kronecker product, and vec and vech operators. This results in a smaller equation that is easier to solve. It should be noted that the dimension of the resultant equation is the dimension of the Stiefel manifold in question. This means that the equation can be efficiently solved. Section 4 provides information about the two types of numerical experiments we used to evaluate our method. The first experiment was an application to ICA. We demonstrated that the proposed method improved the accuracy when compared with the solution generated by a Jacobi-like method, which is an existing method for the JD problem. The other experiment used a larger problem to show that sequences generated by the proposed Newton's method converge quadratically. Section 5 contains our concluding remarks. In these sections, the Stiefel manifold is endowed with the induced metric from the natural inner product in ambient Euclidean space. In contrast to this, we have derived another formula for the representation matrix of the Hessian in Appendix 5, in which the Stiefel manifold is endowed with the canonical metric.

2 Joint diagonalization problem on the Stiefel manifold

2.1 Joint diagonalization problem

Let A_1, A_2, \dots, A_N be N real $n \times n$ symmetric matrices. We consider the following JD problem on the Stiefel manifold $\text{St}(p, n)$ [17]:

Problem 2.1.

$$\begin{aligned} \text{minimize} \quad & f(Y) = - \sum_{l=1}^N \|\text{diag}(Y^T A_l Y)\|_F^2, \\ \text{subject to} \quad & Y \in \text{St}(p, n), \end{aligned} \tag{2.1}$$

$$\tag{2.2}$$

where $\text{St}(p, n) = \{Y \in \mathbb{R}^{n \times p} \mid Y^T Y = I_p\}$ for $p \leq n$, $\|\cdot\|_F$ denotes the Frobenius norm, and $\text{diag}(\cdot)$ denotes the diagonal part of the matrix.

We wish to apply Newton's method to Problem 2.1, so the Hessian $\text{Hess } f$ of f is fundamental. To derive and analyze the Hessian and other requisites, we first review the geometry of $\text{St}(p, n)$.

2.2 The geometry of the Stiefel manifold

In this subsection, we review the geometry of the Stiefel manifold $\text{St}(p, n)$, as discussed in [1, 9].

The tangent space $T_Y \text{St}(p, n)$ of $\text{St}(p, n)$ at $Y \in \text{St}(p, n)$ is

$$T_Y \text{St}(p, n) = \{ \xi \in \mathbb{R}^{n \times p} \mid \xi^T Y + Y^T \xi = 0 \}. \quad (2.3)$$

In later sections, we will make full use of the equivalent form [9]

$$T_Y \text{St}(p, n) = \{ YB + Y_\perp C \mid B \in \text{Skew}(p), C \in \mathbb{R}^{(n-p) \times p} \}, \quad (2.4)$$

rather than Eq. (2.3), where Y_\perp is an arbitrary $n \times (n - p)$ matrix that satisfies $Y^T Y_\perp = 0$ and $Y_\perp^T Y_\perp = I_{n-p}$, and $\text{Skew}(p)$ denotes the set of all $p \times p$ skew-symmetric matrices. We here note that

$$\dim(\text{St}(p, n)) = \frac{p(p-1)}{2} + p(n-p) = \dim(\text{Skew}(p)) + \dim(\mathbb{R}^{(n-p) \times p}), \quad (2.5)$$

which is an important relation for rewriting Newton's equation into a system of $\dim \text{St}(p, n)$ linear equations.

Because $\text{St}(p, n)$ is a submanifold of the matrix Euclidean space $\mathbb{R}^{n \times p}$, it can be endowed with the Riemannian metric

$$\langle \xi_1, \xi_2 \rangle_Y := \text{tr}(\xi_1^T \xi_2), \quad \xi_1, \xi_2 \in T_Y \text{St}(p, n), \quad (2.6)$$

which is induced from the natural inner product in $\mathbb{R}^{n \times p}$. We view $\text{St}(p, n)$ as a Riemannian submanifold of $\mathbb{R}^{n \times p}$ with the above metric. Under this metric, the orthogonal projection P_Y at Y onto $T_Y \text{St}(p, n)$ is expressed as

$$P_Y(W) = W - Y \text{sym}(Y^T W), \quad Y \in \text{St}(p, n), W \in \mathbb{R}^{n \times p}, \quad (2.7)$$

where $\text{sym}(\cdot)$ denotes the symmetric part of the matrix.

In optimization algorithms on the Euclidean space, the line search is performed after computing the search direction. In Riemannian optimization, the concept of a straight line is replaced with a curve (not necessarily geodesic) on a general Riemannian manifold. A retraction on the manifold in question is needed to implement Riemannian optimization algorithms [1]. It defines an appropriate curve for searching a next iterate point on the manifold. We will use the QR retraction R on the Stiefel manifold $\text{St}(p, n)$ [1] defined as

$$R_Y(\xi) = \text{qf}(Y + \xi), \quad Y \in \text{St}(p, n), \xi \in T_Y \text{St}(p, n), \quad (2.8)$$

where $\text{qf}(\cdot)$ denotes the Q factor of the QR decomposition of the matrix. In other words, if a full-rank matrix $W \in \mathbb{R}^{n \times p}$ is uniquely decomposed into $W = QR$, where $Q \in \text{St}(p, n)$ and R is a $p \times p$ upper triangular matrix with strictly positive diagonal entries, then $\text{qf}(W) = Q$.

2.3 The gradient and the Hessian of the objective function

We now return to Problem 2.1. We need the gradient, $\text{grad } f$, and Hessian, $\text{Hess } f$, of the objective function, f , on $\text{St}(p, n)$ to describe Newton's equation for the problem. Newton's equation is defined at each $Y \in \text{St}(p, n)$ as

$$\text{Hess } f(Y)[\xi] = -\text{grad } f(Y), \quad (2.9)$$

where $\xi \in T_Y \text{St}(p, n)$ is an unknown tangent vector. Expressions for the gradient and Hessian can also be found in [17]. In this subsection, we first briefly described how to compute them.

Let \bar{f} be the function on $\mathbb{R}^{n \times p}$ defined in the same way as the right-hand side of (2.1). We note that f is the restriction of \bar{f} to $\text{St}(p, n)$. Because we view the Stiefel manifold, $\text{St}(p, n)$, as a submanifold of the Euclidean space $\mathbb{R}^{n \times p}$, the gradient $\text{grad } f(Y)$ at $Y \in \text{St}(p, n)$ can be expressed as

$$\text{grad } f(Y) = P_Y(\text{grad } \bar{f}(Y)), \quad (2.10)$$

where $\text{grad } \bar{f}$ is the Euclidean gradient of \bar{f} on $\mathbb{R}^{n \times p}$. Furthermore, the Hessian $\text{Hess } f(Y)$ at $Y \in \text{St}(p, n)$ acts on $\xi \in T_Y \text{St}(p, n)$ as

$$\text{Hess } f(Y)[\xi] = P_Y(D(\text{grad } f)(Y)[\xi]). \quad (2.11)$$

Therefore, we only have to compute $\text{grad } \bar{f}(Y)$ and $D(\text{grad } f)(Y)[\xi]$.

We use the relation

$$\|\text{diag}(Y^T A_l Y)\|_F^2 = \text{tr}(\text{diag}(Y^T A_l Y)^2) = \text{tr}(Y^T A_l Y \text{diag}(Y^T A_l Y)) \quad (2.12)$$

to compute the Frechét derivative $D\bar{f}(Y)[\eta]$ with $\eta \in \mathbb{R}^{n \times p}$ as

$$D\bar{f}(Y)[\eta] = -4 \sum_{l=1}^N \eta^T A_l Y \text{diag}(Y^T A_l Y), \quad (2.13)$$

so that

$$\text{grad } \bar{f}(Y) = -4 \sum_{l=1}^N A_l Y \text{diag}(Y^T A_l Y). \quad (2.14)$$

The gradient $\text{grad } f$ of f on $\text{St}(p, n)$ can be obtained using Eq. (2.10) and (2.14) as

$$\text{grad } f(Y) = -4 \sum_{l=1}^N (A_l Y \text{diag}(Y^T A_l Y) - Y \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))). \quad (2.15)$$

We can also express $D(\text{grad } f)(Y)[\xi]$ as

$$D(\text{grad } f)(Y)[\xi] = P_Y(D(\text{grad } \bar{f})(Y)[\xi]) - \xi \text{sym}(Y^T \text{grad } \bar{f}(Y)) - Y \text{sym}(\xi^T \text{grad } \bar{f}(Y)), \quad (2.16)$$

where $D(\text{grad } \bar{f})(Y)[\xi]$ is easily obtained from Eq. (2.14) as

$$D(\text{grad } \bar{f})(Y)[\xi] = -4 \sum_{l=1}^N (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi)). \quad (2.17)$$

We here note that $\text{diag}(\xi^T A_l Y) = \text{diag}((\xi^T A_l Y)^T) = \text{diag}(Y^T A_l \xi)$. It follows from Eq. (2.11) that

$$\begin{aligned} & \text{Hess } f(Y)[\xi] \\ &= P_Y(D(\text{grad } \bar{f})(Y)[\xi] - \xi \text{sym}(Y^T \text{grad } \bar{f}(Y))) \\ &= -4 \sum_{l=1}^N P_Y (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))) , \end{aligned} \quad (2.18)$$

where we have used $P_Y^2 = P_Y$ and $P_Y(Y \text{sym}(\xi^T \text{grad } \bar{f}(Y))) = 0$ in the first equality.

3 Newton's method for the joint diagonalization problem on the Stiefel manifold

3.1 The Hessian of the objective function as a linear transformation on $\text{Skew}(p) \times \mathbb{R}^{(n-p) \times p}$

In Newton's method for minimizing an objective function, F , on a general Riemannian manifold, M , the search direction $\eta \in T_x M$ at the current point $x \in M$ is computed as a solution to Newton's equation,

$$\text{Hess } F(x)[\eta] = -\text{grad } F(x). \quad (3.1)$$

Because we have already obtained the matrix expressions of $\text{grad } f(Y)$ and $\text{Hess } f(Y)[\xi]$, Newton's equation for Problem 2.1 at $Y \in \text{St}(p, n)$ is

$$\begin{aligned} & -4 \sum_{l=1}^N P_Y (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))) \\ &= 4 \sum_{l=1}^N (A_l Y \text{diag}(Y^T A_l Y) - Y \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))). \end{aligned} \quad (3.2)$$

This must be solved for $\xi \in T_Y \text{St}(p, n)$, with Y given. Because ξ is in $T_Y \text{St}(p, n)$, ξ must satisfy $\xi^T Y + Y^T \xi = 0$. Eq. (3.2) appears too difficult to solve. This is because Eq. (3.2) is complicated, and ξ is an $n \times p$ matrix with $p(p-1)/2 + p(n-p) < np$ independent variables.

To overcome these difficulties, we wish to obtain the representation matrix of $\text{Hess } f(Y)$ as a linear transformation on $T_Y \text{St}(p, n)$, for an arbitrarily fixed Y . To this end, we identify $T_Y \text{St}(p, n) \simeq \text{Skew}(p) \times \mathbb{R}^{(n-p) \times p}$ as $\mathbb{R}^{p(p-1)/2 + p(n-p)}$ and view ξ as a $(p(p-1)/2 + p(n-p))$ -dimensional vector. This can be done using the form in Eq. (2.4). We arbitrarily fix a Y_\perp that satisfies $Y^T Y_\perp = 0$ and $Y_\perp^T Y_\perp = I_{n-p}$. Such a Y_\perp can be computed by applying the Gram-Schmidt orthonormalization process to $n-p$ linearly independent column vectors of the matrix $I_p - YY^T$ (see Algorithm 3.1). Then, $\xi \in T_Y \text{St}(p, n)$ can be expressed as

$$\xi = YB + Y_\perp C, \quad B \in \text{Skew}(p), \quad C \in \mathbb{R}^{(n-p) \times p}. \quad (3.3)$$

$\text{Hess } f(Y)[\xi] \in T_Y \text{St}(p, n)$ can also be written as

$$\text{Hess } f(Y)[\xi] = YB_H + Y_\perp C_H, \quad B_H \in \text{Skew}(p), \quad C_H \in \mathbb{R}^{(n-p) \times p}. \quad (3.4)$$

Let $Z_l = Y^T A_l Y$, $Z_l^\perp = Y^T A_l Y_\perp$, and $Z_l^{\perp\perp} = Y_\perp^T A_l Y_\perp$. The B_H and C_H in Eq. (3.4) can be written

$$\begin{aligned}
B_H &= Y^T \text{Hess } f(Y)[\xi] \\
&= -4 \sum_{l=1}^N Y^T P_Y (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))) \\
&= -4 \sum_{l=1}^N \text{skew} (Y^T (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y)))) \\
&= -4 \sum_{l=1}^N \text{skew} \left((Z_l B + Z_l^\perp C) \text{diag}(Z_l) + 2Z_l \text{diag}(Z_l B + Z_l^\perp C) - B \text{sym}(Z_l \text{diag}(Z_l)) \right),
\end{aligned} \tag{3.5}$$

and

$$\begin{aligned}
C_H &= Y_\perp^T \text{Hess } f(Y)[\xi] \\
&= -4 \sum_{l=1}^N Y_\perp^T P_Y (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))) \\
&= -4 \sum_{l=1}^N Y_\perp^T (A_l \xi \text{diag}(Y^T A_l Y) + 2A_l Y \text{diag}(Y^T A_l \xi) - \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y))) \\
&= -4 \sum_{l=1}^N \left(((Z_l^\perp)^T B + Z_l^{\perp\perp} C) \text{diag}(Z_l) + 2(Z_l^\perp)^T \text{diag}(Z_l B + Z_l^\perp C) - C \text{sym}(Z_l \text{diag}(Z_l)) \right),
\end{aligned} \tag{3.6}$$

where $\text{skew}(\cdot)$ denotes the skew-symmetric part of the matrix.

3.2 Kronecker product and the vec and vech operators

The vec operator and the Kronecker product are useful for rewriting a matrix equation. They can be used to transform the matrix into an unknown column vector [13, 15]. The vec operator $\text{vec}(\cdot)$ acts on a matrix $W = (w_{ij}) \in \mathbb{R}^{m \times n}$ as

$$\text{vec}(W) = (w_{11}, \dots, w_{m1}, w_{12}, \dots, w_{m2}, \dots, w_{1n}, \dots, w_{mn}). \tag{3.7}$$

That is, $\text{vec}(W)$ is an mn -dimensional column vector obtained by stacking the columns of W one underneath the other. The Kronecker product of $U \in \mathbb{R}^{m \times n}$ and $V \in \mathbb{R}^{p \times q}$ (denoted by $U \otimes V$) is an $mp \times nq$ matrix defined as

$$U \otimes V = \begin{pmatrix} u_{11}V & \cdots & u_{1n}V \\ \vdots & & \vdots \\ u_{m1}V & \cdots & u_{mn}V \end{pmatrix}. \tag{3.8}$$

The following useful properties of these operators are known.

- For $U \in \mathbb{R}^{m \times p}$, $V \in \mathbb{R}^{p \times q}$, and $W \in \mathbb{R}^{q \times n}$,

$$\text{vec}(UVW) = (W^T \otimes U) \text{vec}(V). \tag{3.9}$$

- There exists an $n^2 \times n^2$ permutation matrix T_n such that

$$\text{vec}(W^T) = T_n \text{vec}(W), \quad W \in \mathbb{R}^{n \times n}, \quad (3.10)$$

where T_n is given by $T_n = \sum_{i,j=1}^n E_{ij}^{(n \times n)} \otimes E_{ji}^{(n \times n)}$, and $E_{ij}^{(p \times q)}$ denotes the $p \times q$ matrix that has (i, j) -components equal to 1, and others equal to 0.

Furthermore, we can easily derive the following properties.

- For $W \in \mathbb{R}^{n \times n}$,

$$\text{vec}(\text{sym}(W)) = \frac{1}{2}(I_{n^2} + T_n) \text{vec}(W), \quad \text{vec}(\text{skew}(W)) = \frac{1}{2}(I_{n^2} - T_n) \text{vec}(W). \quad (3.11)$$

- There exists an $n^2 \times n^2$ diagonal matrix Δ_n such that

$$\text{vec}(\text{diag}(W)) = \Delta_n \text{vec}(W), \quad W \in \mathbb{R}^{n \times n}, \quad (3.12)$$

where $\Delta_n = \sum_{i=1}^n E_{ii}^{(n \times n)} \otimes E_{ii}^{(n \times n)}$.

For $C \in \mathbb{R}^{(n-p) \times p}$ in Eq. (3.3), $\text{vec}(C)$ is an appropriate vector expression of C , because all the elements of C are independent variables. On the contrary, for $B \in \text{Skew}(p)$ in Eq. (3.3), $\text{vec}(B)$ contains p zeros stemmed from the diagonal elements of B . These should be removed. In addition, $\text{vec}(B)$ contains duplicates of each independent variable, because the upper triangular part (excluding the diagonal) of B is the negative of the lower triangular part. Therefore, we use the veck operator [11]. The veck operator $\text{veck}(\cdot)$ acts on an $n \times n$ skew-symmetric matrix S as

$$\text{veck}(S) = (s_{21}, \dots, s_{n1}, s_{32}, \dots, s_{n2}, \dots, s_{n,n-1}). \quad (3.13)$$

That is, $\text{veck}(S)$ is an $n(n-1)/2$ -dimensional column vector obtained by stacking the columns of the lower triangular part of S one underneath the other. There exists a matrix $D_n \in \mathbb{R}^{n^2 \times n(n-1)/2}$ that only depends on n (the size of S) such that

$$\text{vec}(S) = D_n \text{veck}(S), \quad (3.14)$$

where D_n is given by $D_n = \sum_{n \geq i > j \geq 1} \left(E_{n(j-1)+i, j(n-(j+1)/2)-n+i}^{(n^2 \times n(n-1)/2)} - E_{n(i-1)+j, j(n-(j+1)/2)-n+i}^{(n^2 \times n(n-1)/2)} \right)$. We here note that Eq. (3.14) is valid only if S is skew-symmetric. Because each column of D_n has a 1 and a -1 , and because each row of D_n has at most one nonzero element, we have $D_n^T D_n = 2I_{n(n-1)/2}$. It follows that

$$\text{veck}(S) = \frac{1}{2} D_n^T \text{vec}(S). \quad (3.15)$$

3.3 Representation matrix of the Hessian and Newton's equation

We have obtained all the requisites in the previous subsections. We regard the Hessian $\text{Hess } f(Y)$ at $Y \in \text{St}(p, n)$ as a linear transformation H on $\mathbb{R}^{p(p-1)/2 + p(n-p)}$ that transforms a $(p(p-1)/2 + p(n-p))$ -vector, $\begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix}$, into $\begin{pmatrix} \text{veck}(B_H) \\ \text{vec}(C_H) \end{pmatrix}$, where $\xi = YB + Y_\perp C$ and $\text{Hess } f(Y)[\xi] = YB_H + Y_\perp C_H$. Our goal is to obtain the representation matrix H_A of H .

From Eq. (3.5) and (3.6), $\text{veck}(B_H)$ and $\text{vec}(C_H)$ are calculated using

$$\begin{aligned}
& \text{veck}(B_H) \\
&= \frac{1}{2} D_p^T \text{vec} \left(-4 \sum_{l=1}^N \text{skew} \left((Z_l B + Z_l^\perp C) \text{diag}(Z_l) + 2Z_l \text{diag}(Z_l B + Z_l^\perp C) - B \text{sym}(Z_l \text{diag}(Z_l)) \right) \right) \\
&= -2 D_p^T \sum_{l=1}^N \frac{1}{2} (I_{p^2} - T_p) \text{vec} \left((Z_l B + Z_l^\perp C) \text{diag}(Z_l) + 2Z_l \text{diag}(Z_l B + Z_l^\perp C) - B \text{sym}(Z_l \text{diag}(Z_l)) \right) \\
&= -D_p^T (I_{p^2} - T_p) \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes Z_l) \text{vec}(B) + (\text{diag}(Z_l) \otimes Z_l^\perp) \text{vec}(C) \right. \\
&\quad \left. + 2(I_p \otimes Z_l) \text{vec}(\text{diag}(Z_l B + Z_l^\perp C)) - (\text{sym}(Z_l \text{diag}(Z_l)) \otimes I_p) \text{vec}(B) \right) \\
&= -D_p^T (I_{p^2} - T_p) \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes Z_l + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l) - \text{sym}(Z_l \text{diag}(Z_l)) \otimes I_p) D_p \text{veck}(B) \right. \\
&\quad \left. + (\text{diag}(Z_l) \otimes Z_l^\perp + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l^\perp)) \text{vec}(C) \right), \tag{3.16}
\end{aligned}$$

and

$$\begin{aligned}
& \text{vec}(C_H) \\
&= \text{vec} \left(-4 \sum_{l=1}^N \left(((Z_l^\perp)^T B + Z_l^{\perp\perp} C) \text{diag}(Z_l) + 2(Z_l^\perp)^T \text{diag}(Z_l B + Z_l^\perp C) - C \text{sym}(Z_l \text{diag}(Z_l)) \right) \right) \\
&= -4 \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes (Z_l^\perp)^T) \text{vec}(B) + (\text{diag}(Z_l) \otimes Z_l^{\perp\perp}) \text{vec}(C) \right. \\
&\quad \left. + 2(I_p \otimes (Z_l^\perp)^T) \text{vec}(\text{diag}(Z_l B + Z_l^\perp C)) - (\text{sym}(Z_l \text{diag}(Z_l)) \otimes I_{n-p}) \text{vec}(C) \right) \\
&= -4 \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes (Z_l^\perp)^T + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l)) D_p \text{veck}(B) \right. \\
&\quad \left. + (\text{diag}(Z_l) \otimes Z_l^{\perp\perp} + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l^\perp) - \text{sym}(Z_l \text{diag}(Z_l)) \otimes I_{n-p}) \text{vec}(C) \right). \tag{3.17}
\end{aligned}$$

Therefore, we obtain the linear relation

$$\begin{pmatrix} \text{veck}(B_H) \\ \text{vec}(C_H) \end{pmatrix} = H_A \begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix}. \tag{3.18}$$

Here, the representation matrix of H , H_A , is given by

$$H_A = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \tag{3.19}$$

where

$$H_{11} = -D_p^T (I_{p^2} - T_p) \sum_{l=1}^N (\text{diag}(Z_l) \otimes Z_l + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l) - \text{sym}(Z_l \text{diag}(Z_l)) \otimes I_p) D_p, \tag{3.20}$$

$$H_{12} = -D_p^T(I_{p^2} - T_p) \sum_{l=1}^N \left(\text{diag}(Z_l) \otimes Z_l^\perp + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l^\perp) \right), \quad (3.21)$$

$$H_{21} = -4 \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes (Z_l^\perp)^T + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l)) D_p, \quad (3.22)$$

and

$$H_{22} = -4 \sum_{l=1}^N \left(\text{diag}(Z_l) \otimes Z_l^{\perp\perp} + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l^\perp) - \text{sym}(Z_l \text{diag}(Z_l)) \otimes I_{n-p} \right). \quad (3.23)$$

Thus, Newton's equation, $\text{Hess } f(Y)[\xi] = -\text{grad } f(Y)$, can be solved using the following method. We first note that Newton's equation is equivalent to

$$\begin{cases} Y^T \text{Hess } f(Y)[\xi] = -Y^T \text{grad } f(Y), \\ Y_\perp^T \text{Hess } f(Y)[\xi] = -Y_\perp^T \text{grad } f(Y). \end{cases} \quad (3.24)$$

Applying the veck operator to first equation of Eq. (3.24), and the vec operator to the second, yields

$$H_A \begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix} = - \begin{pmatrix} \text{veck}(Y^T \text{grad } f(Y)) \\ \text{vec}(Y_\perp^T \text{grad } f(Y)) \end{pmatrix}, \quad (3.25)$$

where $\xi = YB + Y_\perp C$ with $B \in \text{Skew}(p)$ and $C \in \mathbb{R}^{(n-p) \times p}$. If H_A is invertible, we can solve Eq. (3.25) using

$$\begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix} = -H_A^{-1} \begin{pmatrix} \text{veck}(Y^T \text{grad } f(Y)) \\ \text{vec}(Y_\perp^T \text{grad } f(Y)) \end{pmatrix}. \quad (3.26)$$

After we have obtained $\text{vec}(B) = D_p \text{veck}(B)$ and $\text{vec}(C)$, we can easily reshape $B \in \text{Skew}(p)$ and $C \in \mathbb{R}^{(n-p) \times p}$. Therefore, we have calculated the solution $\xi = YB + Y_\perp C$.

3.4 Newton's method

When implementing Newton's equation, we must compute the matrix H_A in Eq. (3.19). If the block matrices $H_{11}, H_{12}, H_{21}, H_{22}$ of H_A have some relationships, we may reduce the computational cost of computing H_A .

The Hessian, $\text{Hess } F(Y)$, is symmetric with respect to the metric $\langle \cdot, \cdot \rangle_Y$. However, the representation matrix H_A is not symmetric. This is because, for $\xi = YB_1 + Y_\perp C_1$ and $\eta = YB_2 + Y_\perp C_2$, we have

$$\langle \xi, \eta \rangle_Y = \text{tr}(B_1^T B_2) + \text{tr}(C_1^T C_2) = 2 \text{veck}(B_1)^T \text{veck}(B_2) + \text{vec}(C_1)^T \text{vec}(C_2), \quad (3.27)$$

so that the independent coordinates of B_1 and B_2 are counted twice. If we endowed $\text{St}(p, n)$ with the canonical metric [9], the representation matrix would be symmetric (see Appendix A for more details).

Although the representation matrix (H_A) with the induced metric is not symmetric, it does satisfy the following. We have

$$\langle \text{Hess } f(Y)[\xi], \eta \rangle = \langle \text{Hess } f(Y)[\eta], \xi \rangle, \quad \xi, \eta \in T_Y \text{St}(p, n), \quad (3.28)$$

because $\text{Hess } f(Y)$ is symmetric with respect to the induced metric. Let $\xi = YB_1 + Y_\perp C_1$ and $\eta = YB_2 + Y_\perp C_2$. It follows from Eq. (3.18) and (3.27) that

$$\begin{aligned} & \left(H_A \begin{pmatrix} \text{veck}(B_1) \\ \text{vec}(C_1) \end{pmatrix} \right)^T \begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix} \begin{pmatrix} \text{veck}(B_2) \\ \text{vec}(C_2) \end{pmatrix} \\ &= \left(H_A \begin{pmatrix} \text{veck}(B_2) \\ \text{vec}(C_2) \end{pmatrix} \right)^T \begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix} \begin{pmatrix} \text{veck}(B_1) \\ \text{vec}(C_1) \end{pmatrix}. \end{aligned} \quad (3.29)$$

Now

$$\begin{pmatrix} \text{veck}(B_1) \\ \text{vec}(C_1) \end{pmatrix}^T \left(H_A^T \begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix} - \begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix} H_A \right) \begin{pmatrix} \text{veck}(B_2) \\ \text{vec}(C_2) \end{pmatrix} = 0. \quad (3.30)$$

We have

$$\begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix} H_A = H_A^T \begin{pmatrix} 2I_{p(p-1)/2} & 0 \\ 0 & I_{p(n-p)} \end{pmatrix}, \quad (3.31)$$

because $\begin{pmatrix} \text{veck}(B_1) \\ \text{vec}(C_1) \end{pmatrix}$ and $\begin{pmatrix} \text{veck}(B_2) \\ \text{vec}(C_2) \end{pmatrix}$ can be arbitrary $(p(p-1)/2 + p(n-p))$ -dimensional vectors. We can rewrite Eq. (3.31) using the block matrices of H_A as

$$\begin{pmatrix} 2H_{11} & 2H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} 2H_{11}^T & H_{21}^T \\ 2H_{12}^T & H_{22}^T \end{pmatrix}. \quad (3.32)$$

Therefore, the block matrices satisfy

$$H_{11} = H_{11}^T, \quad H_{21} = 2H_{12}^T, \quad H_{22} = H_{22}^T. \quad (3.33)$$

Thus, after we have computed H_{12} , H_{21} can be computed from $H_{21} = 2H_{12}^T$. We do not have to use Eq. (3.22). In fact, Eq. (3.33) can also be shown directly from the formula $D_p^T = -D_p^T T_p$ (see Appendix A).

We proceed to Newton's method for Problem 2.1. Algorithm 3.1 describes how to compute $Y_\perp \in \text{St}(n-p, n)$ that satisfies $Y^T Y_\perp = 0$ and $Y_\perp^T Y_\perp = I_{n-p}$ for a given $Y \in \text{St}(p, n)$. See also [14].

Algorithm 3.1 Method for computing $Y_\perp \in \text{St}(n-p, n)$ for a given $Y \in \text{St}(p, n)$

Input: An orthonormal matrix $Y \in \text{St}(p, n)$.

- 1: Compute $X := I_n - YY^T$.
 - 2: Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ denote the columns of X , that is, $X = (\mathbf{x}_1, \dots, \mathbf{x}_n)$.
 - 3: Set $i = 0$ and $j = 1$.
 - 4: **while** $i < n - p$ **do**
 - 5: **if** $\{\mathbf{x}'_1, \dots, \mathbf{x}'_i, \mathbf{x}_j\}$ is linearly independent **then**
 - 6: $\mathbf{x}'_{i+1} = \mathbf{x}_j$ and $i = i + 1$.
 - 7: **end if**
 - 8: $j = j + 1$.
 - 9: **end while**
 - 10: Set $Z_0 = (\mathbf{x}'_1, \dots, \mathbf{x}'_{n-p})$.
 - 11: Compute $Y_\perp = \text{qf}(Z_0)$.
-

Using Algorithm 3.1 and the QR retraction, we propose Algorithm 3.2 as Newton's method for Problem 2.1. We here note that $D_p^T = -D_p^T T_p$, which is shown in Appendix A.

For the reader's convenience, we also describe Algorithm 3.3. This is Newton's method for the case $n = p$, that is, the case of the orthogonal group. If $n = p$, the relationships $YY^T = I_n$ and $Y_\perp^T = 0$ simplify the algorithm.

4 Application to independent component analysis

4.1 Independent component analysis and the joint diagonalization problem

The simplest ICA model assumes the existence of n independent signals $s_1(t), \dots, s_n(t)$. The observations of n mixtures $x_1(t), \dots, x_n(t)$ are given by the mixing equation

$$\mathbf{x}(t) = A\mathbf{s}(t), \quad (4.1)$$

where $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$, $\mathbf{s}(t) = (s_1(t), \dots, s_n(t))^T$, and A is an $n \times n$ mixing matrix. The problem is to recover the source vector \mathbf{s} , using only the observed data (\mathbf{x}) under the assumption that the entries s_1, \dots, s_n of \mathbf{s} are mutually independent. The problem is formulated as the computation of an $n \times n$ matrix B , which is called a separating matrix, such that

$$\mathbf{z}(t) = B\mathbf{x}(t) \quad (4.2)$$

is an appropriate estimate of the source vector $\mathbf{s}(t)$. In other words, we wish to find B such that the elements z_1, \dots, z_n of \mathbf{z} are mutually independent. See [4] for more details.

The ICA problem is often solved by minimizing an objective function, called a contrast function. One choice for such a function is the JADE (joint approximate diagonalization of eigen-matrices) contrast ϕ , which is the sum of fourth-order cross-cumulants of the elements z_1, z_2, \dots, z_n of \mathbf{z} . We can assume that \mathbf{x} , and therefore \mathbf{z} , are zero-mean random variables because we can subtract the mean $E[\mathbf{x}]$ from \mathbf{x} if needed. The fourth-order cumulants $\mathcal{C}_{ijkl}[\mathbf{z}]$ of zero-mean random variables z_i, z_j, z_k, z_l can be expressed by

$$\mathcal{C}_{ijkl}[\mathbf{z}] = E[z_i z_j z_k z_l] - E[z_i z_j]E[z_k z_l] - E[z_i z_k]E[z_j z_l] - E[z_i z_l]E[z_j z_k]. \quad (4.3)$$

The JADE contrast ϕ of \mathbf{z} is then defined as

$$\phi(\mathbf{z}) = \sum_{\substack{i,j,k,l \\ i \neq j}} (\mathcal{C}_{ijkl}[\mathbf{z}])^2. \quad (4.4)$$

To reformulate the problem as a JD problem, we define cumulant matrices. The cumulant matrix, $Q^{\mathbf{z}}(M)$, associated with a given $n \times n$ matrix $M = (m_{ij})$ is defined to have an (i, j) -th component

$$(Q^{\mathbf{z}}(M))_{ij} = \sum_{k,l=1}^n \mathcal{C}_{ijkl}[\mathbf{z}] m_{kl}. \quad (4.5)$$

If we assume that \mathbf{z} is whitened, the cumulant matrix $Q^{\mathbf{z}}(M)$ can be expressed as

$$Q^{\mathbf{z}}(M) = E[(\mathbf{z}^T M \mathbf{z}) \mathbf{z} \mathbf{z}^T] - \text{tr}(M) I_n - M - M^T. \quad (4.6)$$

Algorithm 3.2 Newton's method for Problem 2.1

- 1: Choose an initial point $Y^{(0)} \in \text{St}(p, n)$.
- 2: **for** $k = 0, 1, 2, \dots$ **do**
- 3: Compute $Y_{\perp}^{(k)}$ that satisfies $(Y^{(k)})^T Y_{\perp}^{(k)} = 0$ and $(Y_{\perp}^{(k)})^T Y_{\perp}^{(k)} = I_{n-p}$, using Algorithm 3.1 with $Y = Y^{(k)}$.
- 4: Compute $Z_l^{(k)} = (Y^{(k)})^T A_l Y^{(k)}$, $Z_l^{\perp(k)} = (Y^{(k)})^T A_l Y_{\perp}^{(k)}$, and $Z_l^{\perp\perp(k)} = (Y_{\perp}^{(k)})^T A_l Y_{\perp}^{(k)}$ for $l = 1, 2, \dots, N$.
- 5: Compute $(Y^{(k)})^T \text{grad } f(Y^{(k)})$ and $(Y_{\perp}^{(k)})^T \text{grad } f(Y^{(k)})$ using

$$(Y^{(k)})^T \text{grad } f(Y^{(k)}) = -4 \text{skew} \left(\sum_{l=1}^N \left(Z_l^{(k)} \text{diag}(Z_l^{(k)}) \right) \right), \quad (3.34)$$

and

$$(Y_{\perp}^{(k)})^T \text{grad } f(Y^{(k)}) = -4 \sum_{l=1}^N \left(\left(Z_l^{\perp(k)} \right)^T \text{diag}(Z_l^{(k)}) \right). \quad (3.35)$$

- 6: Compute the matrices $H_{11}^{(k)}$, $H_{12}^{(k)}$, $H_{21}^{(k)}$, and $H_{22}^{(k)}$ using

$$H_{11}^{(k)} = -2D_p^T \sum_{l=1}^N \left(\text{diag}(Z_l^{(k)}) \otimes Z_l^{(k)} + 2(I_p \otimes Z_l^{(k)}) \Delta_p(I_p \otimes Z_l^{(k)}) \right. \\ \left. - \text{sym}(Z_l^{(k)} \text{diag}(Z_l^{(k)})) \otimes I_p \right) D_p, \quad (3.36)$$

$$H_{12}^{(k)} = -2D_p^T \sum_{l=1}^N \left(\text{diag}(Z_l^{(k)}) \otimes Z_l^{\perp(k)} + 2(I_p \otimes Z_l^{(k)}) \Delta_p(I_p \otimes Z_l^{\perp(k)}) \right), \quad (3.37)$$

$$H_{21}^{(k)} = 2(H_{12}^{(k)})^T, \quad (3.38)$$

$$H_{22}^{(k)} = -4 \sum_{l=1}^N \left(\text{diag}(Z_l^{(k)}) \otimes Z_l^{\perp\perp(k)} + 2(I_p \otimes (Z_l^{\perp(k)})^T) \Delta_p(I_p \otimes Z_l^{\perp(k)}) \right. \\ \left. - \text{sym}(Z_l^{(k)} \text{diag}(Z_l^{(k)})) \otimes I_{n-p} \right), \quad (3.39)$$

where $D_p = \sum_{p \geq i > j \geq 1} \left(E_{p(j-1)+i, j(p-(j+1)/2)-p+i}^{(p^2 \times p(p-1)/2)} - E_{p(i-1)+j, j(p-(j+1)/2)-p+i}^{(p^2 \times p(p-1)/2)} \right)$, and $T_p = \sum_{i,j=1}^p E_{ij}^{(p \times p)} \otimes E_{ji}^{(p \times p)}$.

- 7: Compute $\mathbf{b}^{(k)} \in \mathbb{R}^{p(p-1)/2}$, $\mathbf{c}^{(k)} \in \mathbb{R}^{p(n-p)}$, and $\tilde{\mathbf{b}}^{(k)} \in \mathbb{R}^{p^2}$ using

$$\begin{pmatrix} \mathbf{b}^{(k)} \\ \mathbf{c}^{(k)} \end{pmatrix} = - \begin{pmatrix} H_{11}^{(k)} & H_{12}^{(k)} \\ H_{21}^{(k)} & H_{22}^{(k)} \end{pmatrix}^{-1} \begin{pmatrix} \text{veck}((Y^{(k)})^T \text{grad } f(Y^{(k)})) \\ \text{vec}((Y_{\perp}^{(k)})^T \text{grad } f(Y^{(k)})) \end{pmatrix}, \quad (3.40)$$

and

$$\tilde{\mathbf{b}}^{(k)} = D_p \mathbf{b}^{(k)}. \quad (3.41)$$

- 8: Construct $B^{(k)} \in \text{Skew}(p)$ and $C^{(k)} \in \mathbb{R}^{(n-p) \times p}$ that satisfy $\text{vec}(B^{(k)}) = \tilde{\mathbf{b}}^{(k)}$ and $\text{vec}(C^{(k)}) = \mathbf{c}^{(k)}$.
 - 9: Compute $\xi^{(k)} = Y^{(k)} B^{(k)} + Y_{\perp}^{(k)} C^{(k)}$.
 - 10: Compute the next iterate $Y^{(k+1)} = \text{qf}(Y^{(k)} + \xi^{(k)})$.
 - 11: **end for**
-

Algorithm 3.3 Newton's method for Problem 2.1 on the orthogonal group

- 1: Choose an initial point $Y^{(0)} \in O(n)$.
- 2: **for** $k = 0, 1, 2, \dots$ **do**
- 3: Compute $Z_l^{(k)} = (Y^{(k)})^T A_l Y^{(k)}$ for $l = 1, 2, \dots, N$.
- 4: Compute $\text{grad } f(Y^{(k)})$ using

$$\text{grad } f(Y^{(k)}) = -4 \sum_{l=1}^N \left(A_l Y \text{diag}(Z_l^{(k)}) - Y^{(k)} \text{sym} \left(Z_l^{(k)} \text{diag}(Z_l^{(k)}) \right) \right). \quad (3.42)$$

- 5: Compute the matrix $H_A^{(k)}$ using

$$H_A^{(k)} = -2D_n^T \sum_{l=1}^N \left(\text{diag}(Z_l^{(k)}) \otimes Z_l^{(k)} + 2(I_n \otimes Z_l^{(k)}) \Delta_n(I_n \otimes Z_l^{(k)}) \right. \\ \left. - \text{sym}(Z_l^{(k)} \text{diag}(Z_l^{(k)})) \otimes I_n \right) D_n, \quad (3.43)$$

$$(3.44)$$

where $D_n = \sum_{n \geq i > j \geq 1} \left(E_{n(j-1)+i, j(n-(j+1)/2)-n+i}^{(n^2 \times n(n-1)/2)} - E_{n(i-1)+j, j(n-(j+1)/2)-n+i}^{(n^2 \times n(n-1)/2)} \right)$ and $T_n = \sum_{i,j=1}^n E_{ij}^{(n \times n)} \otimes E_{ji}^{(n \times n)}$.

- 6: Compute $\tilde{\mathbf{b}}^{(k)} \in \mathbb{R}^{n^2}$ using

$$\mathbf{b}^{(k)} = -D_n (H_A^{(k)})^{-1} \text{veck}((Y^{(k)})^T \text{grad } f(Y^{(k)})). \quad (3.45)$$

- 7: Construct $B^{(k)} \in \text{Skew}(p)$ that satisfies $\text{vec}(B^{(k)}) = \tilde{\mathbf{b}}^{(k)}$.
 - 8: Compute $\xi^{(k)} = Y^{(k)} B^{(k)}$.
 - 9: Compute the next iterate $Y^{(k+1)} = \text{qf}(Y^{(k)} + \xi^{(k)})$.
 - 10: **end for**
-

In addition, owing to the assumption of whiteness, we only have to seek a separating matrix B in the orthogonal group $O(n)$. Then, using $\mathbf{z} = B\mathbf{x}$, we can show that [5, 6]

$$\phi(\mathbf{z}) = \sum_{k \leq l} \|\text{off}(Q^{\mathbf{z}}(M_{kl}))\|_F^2 = \sum_{k \leq l} \|\text{off}(BQ^{\mathbf{x}}(M_{kl})B^T)\|_F^2, \quad (4.7)$$

where $\text{off}(\cdot)$ denotes the off-diagonal part of the matrix, and

$$M_{kl} = \begin{cases} E_{kl}^{n \times n} & \text{if } k = l \\ (E_{kl}^{n \times n} + E_{lk}^{n \times n})/\sqrt{2} & \text{if } k < l. \end{cases} \quad (4.8)$$

Therefore, if we define $Q^{\mathbf{x}}(M_{kl}), k \leq l$ as $N := n(n+1)/2$ matrices A_1, \dots, A_N , and define $Y = B^T$, the optimization problem for the JADE contrast is as follows.

Problem 4.1.

$$\text{minimize } g(Y) = \sum_{l=1}^N \|\text{off}(Y^T A_l Y)\|_F^2, \quad (4.9)$$

$$\text{subject to } Y \in O(n). \quad (4.10)$$

Because $Y \in O(n)$, we have

$$\|\text{off}(Y^T A_l Y)\|_F^2 = \|A_l\|_F^2 - \|\text{diag}(Y^T A_l Y)\|_F^2 = -\|\text{diag}(Y^T A_l Y)\|_F^2 + \text{const}. \quad (4.11)$$

Thus, Problem 4.1 is equivalent to Problem 2.1 with $p = n$. In the next section, we apply Algorithm 3.3 to ICA.

4.2 Application to image separation

ICA can be applied to image separation [10]. For simplicity, we will omit the discussion of the process of removing the mean values and whitening, assuming the zero-mean property and whiteness if needed.

We used the three images shown in Fig. 4.1, which were expressed by 256×256 matrices, I_1, I_2, I_3 . We regarded the three images as mutually independent signals using the following



Figure 4.1: Test images as source signals.

method. We let $s_i := \text{vec}(I_i), i = 1, 2, 3$ denote $T(= 256^2)$ -dimensional column vectors, and let $s_i(t)$ denote the t -th element of s_i . We consider that each s_i has T samples. Furthermore, we define the source matrix $S := (s_1, s_2, s_3)^T \in \mathbb{R}^{3 \times T}$. We then mix the source signals

using a mixing matrix $A \in \mathbb{R}^{3 \times 3}$ to obtain $X = AS$ as observed signals. In our experiments, $A = \begin{pmatrix} 0.3873 & 0.4439 & 0.5734 \\ 0.4857 & 0.3472 & 0.6026 \\ 0.1437 & 0.3132 & 0.7527 \end{pmatrix}$. The images of the observed signals, X , are shown in Fig. 4.2. We wished to find a separating matrix B , such that $Z := BX$ is as mutually independent as



Figure 4.2: Mixed images caused by a mixing matrix A .

possible, without using any information from S .

We first computed $n(n+1)/2 = 6$ matrices $A_1 = Q^X(M_{11}), A_2 = Q^X(M_{12}), A_3 = Q^X(M_{22}), A_4 = Q^X(M_{13}), A_5 = Q^X(M_{23}), A_6 = Q^X(M_{33})$, as discussed in the previous subsection. Here, we regard the operation $E[\cdot]$ as the sample mean. All that is left is to jointly diagonalize A_1, A_2, \dots, A_6 . That is, to solve Problem 4.1. Because Newton's method has only a local convergence property, we need an approximate solution of the problem in advance. One way to calculate this approximate solution is to use the Jacobi-like method proposed by Cardoso and Souloumiac [7]. In this experiment, we obtained an approximate solution Y_J using the Jacobi-like method, and then applied Newton's method (Algorithm 3.3) with the initial point Y_J to obtain Y_N . After that, we computed $Z = Y_N^T X$ and estimated the separated images (Fig. 4.3) as J_1, J_2, J_3 , such that $\text{vec}(J_i)$ is the i -th column of the $T \times 3$ matrix Z^T for $i = 1, 2, 3$. Note that, because ICA cannot identify the correct ordering or scaling of the source signals, we have artificially ordered and scaled the estimated signals to obtain Fig. 4.3.



Figure 4.3: Estimated images obtained by a combination of the Jacobi-like and proposed methods.

The advantages of the proposed method can be seen by comparing our solution Y_N with Y_J .

- The value of the objective function g defined by Eq. (4.9):

$$g(Y_J) - g(Y_N) = 5.3291 \times 10^{-15} > 0. \quad (4.12)$$

- Norm of the gradient of the objective function g :

$$\|\text{grad } g(Y_N)\|_{Y_N} = 5.9323 \times 10^{-15}, \|\text{grad } g(Y_J)\|_{Y_J} = 1.1751 \times 10^{-8}, \quad (4.13)$$

$$\|\text{grad } g(Y_J)\|_{Y_J} - \|\text{grad } g(Y_N)\|_{Y_N} = 1.1751 \times 10^{-8} > 0. \quad (4.14)$$

Note that both norms ($\|\cdot\|_{Y_N}$ and $\|\cdot\|_{Y_J}$) are the same Frobenius norm $\|\cdot\|_F$, if we regard the tangent vectors as elements in $\mathbb{R}^{3 \times 3}$. Although $g(Y_N) < g(Y_J)$ in our experiment, this is not enough to say that our proposed method always decreases the value of the objective function because the difference (5.3291×10^{-15}) is too small. We will see this matter further in a larger example problem in the next subsection. However, we can at least observe that the proposed method does not degrade the solution in view of the cost of the objective function. Moreover, it is clear that the proposed method improves the solution in view of the gradient of the objective function. Therefore, the proposed method can improve approximate solutions, so that mixed sources can be separated more clearly.

4.3 Numerical experiments for larger problems

To more intensively investigate the performance of the proposed algorithm, we return to Problem 2.1 and consider the case $n = p = 50, N = 10$. We prepared N randomly chosen $n \times n$ symmetric matrices A_1, A_2, \dots, A_N . In a manner similar to that in the previous subsection, we first applied the Jacobi-like method [7] to obtain an approximate solution Y_J . We then applied the proposed Newton method to obtain Y_N . The results were as follows.

- Value of the objective function f defined by (4.9):

$$f(Y_J) - f(Y_N) = 6.2437 \times 10^{-9} > 0. \quad (4.15)$$

- Norm of the gradient of the objective function f :

$$\|\text{grad } f(Y_J)\|_{Y_J} - \|\text{grad } f(Y_N)\|_{Y_N} = 9.7064 \times 10^{-4} > 0. \quad (4.16)$$

- Orthogonality:

$$\|Y_J' Y_J - I_p\|_F - \|Y_N' Y_N - I_p\|_F = 8.7202 \times 10^{-13} > 0. \quad (4.17)$$

It is obvious that the proposed method improves the accuracy of the approximate solution.

We performed another experiment for $p < n$. In this case, $n = 50, p = 30, N = 10$ and A_1, A_2, \dots, A_N were constructed as follows. We constructed N randomly chosen $n \times n$ diagonal matrices $\Lambda_1, \Lambda_2, \dots, \Lambda_N$, and a randomly chosen $n \times n$ orthogonal matrix P , where the diagonal elements $\lambda_1^{(i)}, \dots, \lambda_n^{(i)}$ of each Λ_i are positive and in descending order. We then computed A_1, A_2, \dots, A_N as $A_i = P \Lambda_i P^T$, $i = 1, 2, \dots, N$. Note that $Y_{\text{opt}} := P I_{n,p}$ is an optimal solution to the problem. We computed an approximate solution $Y_{\text{app}} := \text{qf}(Y_{\text{opt}} + Y_{\text{rand}})$, where Y_{rand} is a randomly chosen $n \times p$ matrix that has elements less than 0.01 (absolute values). With Y_{app} obtained as an initial point, we applied the proposed Newton's method. We compared the accuracy of the resultant solution Y_N (obtained after 5 iterations of the proposed method) with that of Y_{app} .

- Difference between the objective function f and the optimal value $f(Y_{\text{opt}})$:

$$f(Y_{\text{app}}) - f(Y_{\text{opt}}) = 0.1355, \quad f(Y_N) - f(Y_{\text{opt}}) = 1.4211 \times 10^{-14}. \quad (4.18)$$

- Norm of the gradient of the objective function f :

$$\|\text{grad } f(Y_{\text{app}})\|_{Y_{\text{app}}} = 2.4386, \quad \|\text{grad } f(Y_N)\|_{Y_N} = 2.0630 \times 10^{-13}. \quad (4.19)$$

As expected, we can observe from Fig. 4.4 the quadratic convergence of the sequence generated by the proposed method.

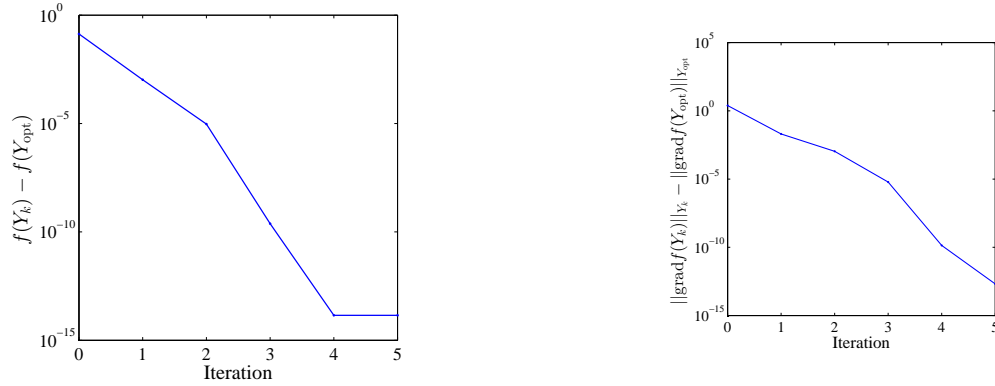


Figure 4.4: Values of the objective function and norms of the gradient of the objective function, obtained with 5 iterations of the proposed method.

5 Concluding remarks

We have considered the joint diagonalization problem on the Stiefel manifold $\text{St}(p, n)$, and have developed Newton's method for the problem. It is difficult to solve Newton's equation, $\text{Hess } f(Y)[\xi] = -\text{grad } f(Y)$, in its original form because we must find an $n \times p$ matrix ξ under the condition $\xi^T Y + Y^T \xi = 0$. To resolve this, we have computed the representation matrix of the Hessian of the objective function, using the Kronecker product and the vec and vech operators. The representation matrix is a $\dim(\text{St}(p, n)) \times \dim(\text{St}(p, n))$ symmetric matrix, and we have succeeded in reducing Newton's equation into the form $Ax = b$ with dimension $\dim(\text{St}(p, n))$, which is less than np . Therefore, the resultant equation can be efficiently computed. With this reduced equation, we have developed a new algorithm for the JD problem.

Furthermore, we have performed numerical experiments to check that the present algorithm is competent for practical applications, and that the algorithm has quadratic convergence. Specifically, we have applied the proposed method to the image separating problem as an example of independent component analysis, and have solved larger problems to more clearly see the performance of our algorithm.

A Newton's equation for Problem 2.1, with respect to the canonical metric

In Section 3, we have endowed the Stiefel manifold $\text{St}(p, n)$ with the induced metric from the natural inner product in $\mathbb{R}^{n \times p}$. In this section, we endow $\text{St}(p, n)$ with the metric g defined by

$$g_Y(\xi, \eta) = \text{tr} \left(\xi^T \left(I_n - \frac{1}{2} Y Y^T \right) \eta \right), \quad \xi, \eta \in T_Y \text{St}(p, n), \quad (\text{A.1})$$

which is called the canonical metric on the Stiefel manifold [9]. If we let $\xi = Y B_1 + Y_\perp C_1$ and $\eta = Y B_2 + Y_\perp C_2$, with $B_1, B_2 \in \text{Skew}(p)$ and $C_1, C_2 \in \mathbb{R}^{(n-p) \times p}$, then

$$g_Y(\xi, \eta) = \text{veck}(B_1)^T \text{veck}(B_2) + \text{vec}(C_1)^T \text{vec}(C_2). \quad (\text{A.2})$$

Thus, the representation matrix of the Hessian with respect to the canonical metric is a symmetric matrix. We shall derive the formula for the representation matrix in a manner similar to that in Section 3.

We note that the gradient and the Hessian of f on $\text{St}(p, n)$ depend on the metric. For clarity, let $\text{grad}^c f$ and $\text{Hess}^c f$ denote the gradient and the Hessian of f with respect to the canonical metric g . Let \bar{f} be an extension of f to $\mathbb{R}^{n \times p}$. According to [9], the gradient $\text{grad}^c f$ and the Hessian quadratic form $g_Y(\text{Hess}^c f(Y)[\xi], \eta)$ are

$$\begin{aligned} \text{grad}^c f(Y) &= \text{grad} \bar{f}(Y) - Y (\text{grad} \bar{f}(Y))^T Y \\ &= -4 \sum_{l=1}^N (A_l Y \text{diag}(Y^T A_l Y) - Y \text{diag}(Y^T A_l Y) Y^T A_l Y), \end{aligned} \quad (\text{A.3})$$

and

$$g_Y(\text{Hess}^c f(Y)[\xi], \eta) = \text{tr} (G_Y(\xi)^T \eta), \quad (\text{A.4})$$

where we have defined G_Y by

$$\begin{aligned} G_Y(\xi) &= D(\text{grad} \bar{f})(Y)[\xi] \\ &\quad + \frac{1}{2} (Y \xi^T \text{grad} \bar{f}(Y) + \text{grad} \bar{f}(Y) \xi^T Y - (I_n - Y Y^T) \xi (Y^T \text{grad} \bar{f}(Y) + \text{grad} \bar{f}(Y)^T Y)) \\ &= -4 \sum_{l=1}^N (A_l \xi \text{diag}(Y^T A_l Y) + 2 A_l Y \text{diag}(Y^T A_l \xi) \\ &\quad + \frac{1}{2} (Y \xi^T A_l Y \text{diag}(Y^T A_l Y) + A_l Y \text{diag}(Y^T A_l Y) \xi^T Y \\ &\quad - (I_n - Y Y^T) \xi \text{sym}(Y^T A_l Y \text{diag}(Y^T A_l Y)))) \end{aligned} \quad (\text{A.5})$$

and $\text{grad} \bar{f}$ is the standard Euclidean gradient of \bar{f} (as in Section 3). Note that $G_Y(\xi)$ is not $\text{Hess} f(Y)[\xi]$, because the metric g is not the one induced from the natural inner product. Using the induced metric $\langle \cdot, \cdot \rangle$, and the orthogonal projection (2.7) with respect to $\langle \cdot, \cdot \rangle$, we obtain

$$\begin{aligned} g_Y(\text{Hess}^c f(Y)[\xi], \eta) &= \text{tr} (G_Y(\xi)^T \eta) \\ &= \langle P_Y(G_Y(\xi)), \eta \rangle_Y \end{aligned}$$

$$\begin{aligned}
&= \text{tr} \left(\left(Y(Y^T P_Y(G_Y(\xi))) + Y_\perp(Y_\perp^T P_Y(G_Y(\xi))) \right)^T (Y(Y^T \eta) + Y_\perp(Y_\perp^T \eta)) \right) \\
&= \text{tr} \left(\left(Y^T P_Y(G_Y(\xi)) \right)^T (Y^T \eta) + (Y_\perp^T P_Y(G_Y(\xi)))^T (Y_\perp^T \eta) \right) \\
&= \text{tr} \left(\frac{1}{2} (2Y^T P_Y(G_Y(\xi)))^T (Y^T \eta) + (Y_\perp^T P_Y(G_Y(\xi)))^T (Y_\perp^T \eta) \right) \\
&= g_Y (2Y Y^T P_Y(G_Y(\xi)) + Y_\perp Y_\perp^T P_Y(G_Y(\xi)), \eta) \\
&= g_Y ((I_n + Y Y^T) P_Y(G_Y(\xi)), \eta), \tag{A.6}
\end{aligned}$$

where we have used $Y_\perp Y_\perp^T = I_n - Y Y^T$, and where $Y_\perp \in \text{St}(n-p, n)$ is an arbitrary matrix which satisfies $Y^T Y_\perp = 0$ and $Y_\perp^T Y_\perp = I_{n-p}$. Because $(I_n + Y Y^T) P_Y(G_Y(\xi))$ is also a tangent vector at $Y \in \text{St}(p, n)$, we have

$$\text{Hess } f(Y)[\xi] = (I_n + Y Y^T) P_Y(G_Y(\xi)). \tag{A.7}$$

Here, $\xi = Y B + Y_\perp C$ and $\text{Hess } f(Y)[\xi] = Y B_H^c + Y_\perp C_H^c$, where $B, B_H^c \in \text{Skew}(p)$ and $C, C_H^c \in \mathbb{R}^{(n-p) \times p}$. Let $Z_l = Y^T A_l Y$, $Z_l^\perp = Y^T A_l Y_\perp$, and $Z_l^{\perp\perp} = Y_\perp^T A_l Y_\perp$. B_H^c and C_H^c are

$$\begin{aligned}
B_H^c &= Y^T \text{Hess } f(Y)[\xi] \\
&= 2 \text{skew}(Y^T G_Y(\xi)) \\
&= -8 \text{skew} \left(\sum_{l=1}^N \left((Z_l B + Z_l^\perp C) \text{diag}(Z_l) + 2Z_l \text{diag}(Z_l B + Z_l^\perp C) \right. \right. \\
&\quad \left. \left. + \frac{1}{2} \left((-B Z_l + C^T (Z_l^\perp)^T) \text{diag}(Z_l) - Z_l \text{diag}(Z_l B) \right) \right) \right), \tag{A.8}
\end{aligned}$$

and

$$\begin{aligned}
C_H^c &= Y_\perp^T \text{Hess } f(Y)[\xi] \\
&= Y_\perp^T G_Y(\xi) \\
&= -4 \sum_{l=1}^N \left(((Z_l^\perp)^T B + Z_l^{\perp\perp} C) \text{diag}(Z_l) + 2(Z_l^\perp)^T \text{diag}(Z_l B + Z_l^\perp C) \right. \\
&\quad \left. - \frac{1}{2} (Z_l^\perp)^T \text{diag}(Z_l) B - C \text{sym}(Z_l \text{diag}(Z_l)) \right). \tag{A.9}
\end{aligned}$$

Therefore, we obtain

$$\begin{pmatrix} \text{veck}(B_H^c) \\ \text{vec}(C_H^c) \end{pmatrix} = H_A^c \begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix}, \tag{A.10}$$

where the representation matrix H_A^c is given by $\begin{pmatrix} H_{11}^c & H_{12}^c \\ H_{21}^c & H_{22}^c \end{pmatrix}$ with

$$\begin{aligned}
H_{11}^c &= -2D_p^T(I_{p^2} - T_p) \\
&\times \sum_{l=1}^N (\text{diag}(Z_l) \otimes Z_l + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l) - (\text{sym}(Z_l \text{diag}(Z_l)) \otimes I_p)) D_p, \tag{A.11}
\end{aligned}$$

$$H_{12}^c = -2D_p^T(I_{p^2} - T_p) \sum_{l=1}^N \left(\text{diag}(Z_l) \otimes Z_l^\perp + 2(I_p \otimes Z_l) \Delta_p(I_p \otimes Z_l^\perp) - \frac{1}{2} I_p \otimes \text{diag}(Z_l) Z_l^\perp \right), \quad (\text{A.12})$$

$$H_{21}^c = -4 \sum_{l=1}^N \left((\text{diag}(Z_l) \otimes (Z_l^\perp)^T + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l) - \frac{1}{2} I_p \otimes (Z_l^\perp)^T \text{diag}(Z_l)) D_p, \quad (\text{A.13})$$

and

$$H_{22}^c = -4 \sum_{l=1}^N \left(\text{diag}(Z_l) \otimes Z_l^{\perp\perp} + 2(I_p \otimes (Z_l^\perp)^T) \Delta_p(I_p \otimes Z_l^\perp) - \text{sym}(Z_l \text{diag}(Z_l)) \otimes I_{n-p} \right). \quad (\text{A.14})$$

Therefore, the solution ξ to Newton's equation,

$$\text{Hess}^c f(Y)[\xi] = -\text{grad}^c f(Y), \quad (\text{A.15})$$

is $\xi = YB + Y_\perp C$, where B and C satisfy

$$\begin{pmatrix} \text{veck}(B) \\ \text{vec}(C) \end{pmatrix} = -(H_A^c)^{-1} \begin{pmatrix} \text{veck}(Y^T \text{grad}^c f(Y)) \\ \text{vec}(Y_\perp^T \text{grad}^c f(Y)) \end{pmatrix}. \quad (\text{A.16})$$

We here note that H_A^c should be symmetric, so that

$$(H_{11}^c)^T = H_{11}^c, \quad (H_{12}^c)^T = H_{21}^c, \quad (H_{22}^c)^T = H_{22}^c. \quad (\text{A.17})$$

We can also directly derive Eq. (3.33) and (A.17) from the expressions of the block matrices of H_A and H_A^c using

$$(U \otimes V)^T = U^T \otimes V^T, \quad U \in \mathbb{R}^{m \times n}, \quad V \in \mathbb{R}^{p \times q}, \quad (\text{A.18})$$

and

$$D_p^T = -D_p^T T_p. \quad (\text{A.19})$$

We shall now derive Eq. (A.19). For any natural number $i \leq p(p-1)/2$, there exist unique integers q and r such that $i = (p-1) + \dots + (p-q) + r$, $0 \leq q \leq p-2$, and $1 \leq r \leq p-q-1$. For these q and r , we define ϕ_1 and ϕ_2 as $\phi_1(i) = q + r + 1$ and $\phi_2(i) = q + 1$. For a $p \times p$ matrix U , each element of $D_p^T \text{vec}(U)$ is the sum of the $(\phi_1(i), \phi_2(i))$ elements and $-(\phi_2(i), \phi_1(i))$ elements of U . Thus, for a $p \times p$ skew-symmetric matrix V and a $p \times p$ symmetric matrix W , we obtain

$$\frac{1}{2} D_p^T \text{vec}(V) = \text{veck}(V), \quad (\text{A.20})$$

and

$$\frac{1}{2} D_p^T \text{vec}(W) = 0. \quad (\text{A.21})$$

On the other hand, it follows from Eq. (3.10) that

$$-\frac{1}{2} D_p^T T_p \text{vec}(V) = -\frac{1}{2} D_p^T \text{vec}(V^T) = \frac{1}{2} D_p^T \text{vec}(V) = \text{veck}(V) \quad (\text{A.22})$$

and

$$-\frac{1}{2} D_p^T T_p \text{vec}(W) = -\frac{1}{2} D_p^T \text{vec}(W^T) = -\frac{1}{2} D_p^T \text{vec}(W) = 0. \quad (\text{A.23})$$

Therefore, for an arbitrary matrix U , it holds that

$$\begin{aligned} -\frac{1}{2}D_p^T T_p \text{vec}(U) &= -\frac{1}{2}D_p^T T_p \text{vec}(\text{sym}(U) + \text{skew}(U)) = \text{veck}(\text{skew}(U)) \\ &= \frac{1}{2}D_p^T \text{vec}(\text{sym}(U) + \text{skew}(U)) = \frac{1}{2}D_p^T \text{vec}(U). \end{aligned} \quad (\text{A.24})$$

Because U is arbitrary, $\text{vec}(U)$ is also an arbitrary p^2 -dimensional column vector, so that $-D_p^T T_p = D_p^T$. This completes the proof.

Acknowledgment

The author appreciates the JSPS Research Fellowship for Young Scientists.

References

- [1] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [2] B. AFSARI AND P. S. KRISHNAPRASAD, *Some gradient based joint diagonalization methods for ica*, in *Independent Component Analysis and Blind Signal Separation*, Springer, 2004, pp. 437–444.
- [3] A. BUNSE-GERSTNER, R. BYERS, AND V. MEHRMANN, *Numerical methods for simultaneous diagonalization*, *SIAM Journal on Matrix Analysis and Applications*, 14 (1993), pp. 927–949.
- [4] J.-F. CARDOSO, *Blind signal separation: statistical principles*, *Proceedings of the IEEE*, 86 (1998), pp. 2009–2025.
- [5] J.-F. CARDOSO, *High-order contrasts for independent component analysis*, *Neural computation*, 11 (1999), pp. 157–192.
- [6] J.-F. CARDOSO AND A. SOULOUMIAC, *Blind beamforming for non-gaussian signals*, in *Radar and Signal Processing*, *IEE Proceedings F*, vol. 140, IET, 1993, pp. 362–370.
- [7] J.-F. CARDOSO AND A. SOULOUMIAC, *Jacobi angles for simultaneous diagonalization*, *SIAM Journal on Matrix Analysis and Applications*, 17 (1996), pp. 161–164.
- [8] A. CICHOCKI, S.-I. AMARI, ET AL., *Adaptive blind signal and image processing*, John Wiley Chichester, 2002.
- [9] A. EDELMAN, T. A. ARIAS, AND S. T. SMITH, *The geometry of algorithms with orthogonality constraints*, *SIAM journal on Matrix Analysis and Applications*, 20 (1998), pp. 303–353.
- [10] H. FARID AND E. H. ADELSON, *Separating reflections from images by use of independent component analysis*, *JOSA A*, 16 (1999), pp. 2136–2145.
- [11] E. W. GRAFAREND, *Linear and nonlinear models: fixed effects, random effects, and mixed models*, De Gruyter, 2006.

- [12] U. HELMKE, J. B. MOORE, AND W. GERMANY, *Optimization and dynamical systems*, (1994).
- [13] H. NEUDECKER AND J. R. MAGNUS, *Matrix differential calculus with applications in statistics and econometrics*, 1999.
- [14] H. SATO AND T. IWAI, *Optimization algorithms on the grassmann manifold with application to matrix eigenvalue problems*, Japan Journal of Industrial and Applied Mathematics, (to appear).
- [15] J. R. SCHOTT, *Matrix analysis for statistics*, (2005).
- [16] S. T. SMITH, *Optimization techniques on riemannian manifolds*, Fields institute communications, 3 (1994), pp. 113–135.
- [17] F. THEIS, T. P. CASON, AND P.-A. ABSIL, *Soft dimension reduction for ica by joint diagonalization on the stiefel manifold*, in Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation, ICA '09, Berlin, Heidelberg, 2009, Springer-Verlag, pp. 354–361.
- [18] M. WAX AND J. SHEINVALD, *A least-squares approach to joint diagonalization*, Signal Processing Letters, IEEE, 4 (1997), pp. 52–53.
- [19] I. YAMADA AND T. EZAKI, *An orthogonal matrix optimization by dual cayley parametrization technique*, in 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), 2003.
- [20] A. YEREDOR, *Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation*, Signal Processing, IEEE Transactions on, 50 (2002), pp. 1545–1553.